

<http://autonom.edu.pl>

Dr inż. Jarosław Krzywański

Mgr inż. Maciej Węgrzyn

Wykorzystanie sztucznych sieci neuronowych dla celów modelowania rzeczywistości a teoria systemów autonomicznych M. Mazura

Streszczenie

W pracy zaprezentowano podstawowe zagadnienia z zakresu techniki sieci neuronowych. Opisano budowę oraz pracę sieci ze szczególnym uwzględnieniem procesu uczenia sieci. Podano podstawowe wymogi niezbędne dla prawidłowego funkcjonowania sztucznych sieci neuronowych oraz zaprezentowano wybrane ich praktyczne zastosowania.

Ponadto podano analogie sieci neuronowych z budową korelatora systemu autonomicznego M. Mazura.

Wstęp

Sztuczne sieci neuronowe stanowią alternatywną metodę symulacji zdolną do odwzorowywania nadzwyczaj złożonych funkcji a nawet do modelowania słabo zdefiniowanych procesów fizycznych [1].

Podstawową ich zaletą jest zdolność do zamodelowania skomplikowanych, nieliniowych zagadnień na drodze prezentacji przykładowych danych wejściowych, ewentualnie zestawów danych wejściowych i wyjściowych. Posiadają one zdolność do obsługi dużych i skomplikowanych systemów z wieloma wzajemnie powiązаныmi wielkościami. Cechę tę zawdzięczają umiejętności „uczenia się” zależności istniejących między parametrami wejściowymi i wyjściowymi. Prezentowane sieci dane mogą zawierać sygnały o charakterze zakłóceń a nawet sygnały nadmiarowe [2]. Sieci „potrafią” ignorować dane nadmiarowe oraz te, których wpływ na badane zjawisko jest pomijalnie mały, „koncentrując” się na wielkościach wejściowych o decydującym znaczeniu dla opisu modelowanego procesu [3].

Powyższe zadecydowało o szerokim ich wykorzystaniu w wielu gałęziach nauki i techniki. Dotyczy to w szczególności zagadnień optymalizacji, nadzoru, robotyki, prognozowania nawet obróbki sygnałów. Jako przykłady wymienić tu można: modelowanie i sterowanie

obiektów dynamicznych, predykcja obciążeń systemu elektroenergetycznego, rozpoznawanie wzorców, kompresja danych [1, 2, 3, 4, 5, 6, 7].

Szczególnie użyteczne są w obszarze zagadnień związanych z modelowaniem. Modele oparte o technikę obliczeń wykorzystujących sztuczne sieci neuronowe mogą być traktowane jako jedne z metod predykcji oraz analizy inżynierskiej [2, 3].

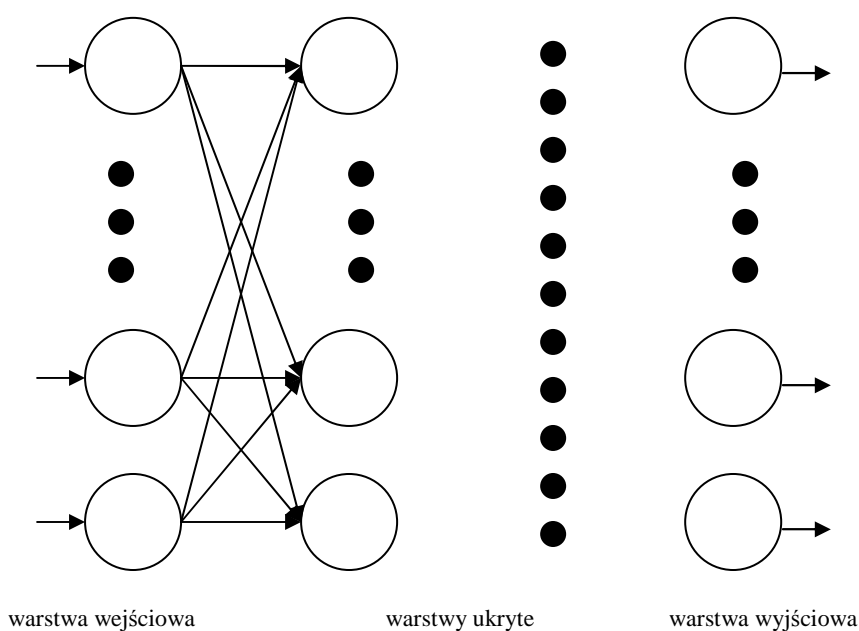
W artykule wykorzystano sztuczną sieć neuronową dla celów oceny wpływu parametrów eksploatacyjnych kotła CFB o mocy 670 MW na wartość efektywnego współczynnika przejmowania ciepła w cyrkulacyjnej warstwie fluidalnej.

Wartość współczynnika przejmowania ciepła oraz jego zależność od warunków eksploatacyjnych została wcześniej określona numerycznie [8] i potwierdzona eksperymentalnie [8].

Stworzenie modelu wykorzystującego zalety sieci neuronowej umożliwiło prowadzenie analiz pozwalających rozpoznać wpływ zmiennych warunków eksploatacyjnych kotła na wartości współczynnika przejmowania ciepła w cyrkulacyjnej warstwie fluidalnej.

Struktura sieci

Sztuczne sieci neuronowe naśladują proces uczenia mózgu człowieka [1, 3]. Sieć stanowi grupa wzajemnie połączonych neuronów. Rysunek nr 1 przedstawia uproszczony schemat wielowarstwowej sieci, w skład której wchodzi trzy rodzaje warstw neuronów: warstwa wejściowa, warstwy ukryte oraz warstwa wyjściowa.



Rys. 1. Przykładowa sieć neuronowa

Sieć neuronowa gromadzi i przechowuje informacje w postaci „siły” wzajemnych połączeń neuronów, zwanych wagami. Z uwagi (szczególnie występuje to dla dużych sieci neuronowych) na brak praktycznej możliwości prawidłowego ustawienia wag, konieczne jest wykorzystanie zdolności tzw. „uczenia się sieci”, celem ustawienia wag [2]. Proces taki odbywa się w fazie nauki. Jedną z metod realizacji fazy nauki jest tzw. uczenie nadzorowane, w czasie którego każda z danych wejściowych oraz odpowiadający jej wzorzec są wielokrotnie prezentowane sieci. Różnica pomiędzy wartością wzorca i odpowiadającego mu, wygenerowanego przez sieć, sygnału wyjściowego, stanowi miarę stopnia modyfikacji wag.

Tutaj należy zwrócić uwagę czytelnika na analogię zachodzącą pomiędzy budowa sieci i jej działaniem a budowa korelatora systemu autonomicznego¹. W korelatorze potencjały wejściowe gromadzone są w wyznaczonych punktach zupełnie jak w pierwszej warstwie sieci. Wielkość przekazywanego sygnału wg Mazura zależy od przewodności dróg korelacyjnych - w sieci przewodność jest zastąpiona pojęciem wagi, ale wielkość wagi oznacza wielkość sygnału, a to samo znaczenie ma wielkość przewodności dróg korelacyjnych. Zmiana przewodności jest wynikiem wielokrotnego przepływu sygnału - w pełni analogicznie do sieci. Nie używane drogi przewodności tracą przewodność też analogicznie jak w sieci. Różnica polega na tym, że u Mazura przewodność zmienia się płynnie a w sieci skokowo, gdyż u Mazura korelator ma strukturę ciągłą /continuum/, a w sieci struktura jest dyskretna, pofragmentowana /granulatum/ [10]. Warto również zauważyć, że wyszkolona sieć może wyłączać niektóre swoje fragmenty - tak np. jak człowiek, który ucząc się jeździć na rowerze najpierw nie jest zdolny myśleć o czym innym niż utrzymywanie równowagi, a po wytrenowaniu w ogóle o tym nie musi myśleć, wykorzystując umysł do myślenia /na przykład o pięknie krajobrazu/.

Najbardziej popularnym i skutecznym algorytmem uczenia wielowarstwowej sieci neuronowej jest metoda wstecznej propagacji błędów (ang. BackPropagation - BP). Modyfikacjami algorytmu BP są: momentowa metoda wstecznej propagacji błędów, metody ze zmiennymi współczynnikami uczenia i momentu oraz metody uwzględniające alternatywną postać funkcji błędu [1].

Aby uzyskać dokładny model dla całej przestrzeni zmian wartości parametrów wejściowych sieć musi posiadać umiejętność uogólniania. W tym celu w procesie uczenia

¹ M. Mazur, Cybernetyka i charakter, Warszawa 1999, str. 256.

należy uwzględnić dostatecznie dużą liczbę danych uczących, pokrywających swymi wartościami cały zakres zmienności [11].

Sieć neuronowa wyuczona na wąskim zakresie danych wejściowych może nie mieć zdolności do generowania prawidłowych odpowiedzi. Z drugiej jednak strony, sieć wyuczona na zbyt dużym zakresie może utracić zdolności do koncentracji na rzeczywistych zależnościach a stąd dokładność modelu również może ulec pogorszeniu [3].

Poprawę zdolności i dokładności przewidywania sieci neuronowej można uzyskać stosując odpowiednią liczbę neuronów ukrytych oraz poprzez dobór właściwej funkcji aktywacji. Podczas modelowania przy użyciu sieci neuronowych funkcję aktywacji z reguły wybiera twórca sieci a wagi są dopasowywane w fazie nauki sieci.

W wielowarstwowych sieciach neuronowych uczonych metodą BP najczęściej stosowanymi funkcjami aktywacji, są funkcje [3, 4, 5]:

sigmoidalna:

$$f(x) = \frac{1}{1 + e^{-\beta x}},$$

oraz tangensoidalna:

$$f(x) = \frac{e^{\alpha x} - e^{-\alpha x}}{e^{\alpha x} + e^{-\alpha x}}.$$

Jednym z kryteriów zakończenia fazy nauki jest wartość błędu średniokwadratowego na wyjściu sieci.

Po przejściu fazy uczenia sieć poddawana jest testom przy użyciu zestawu próbek losowo wybranych spośród zakresu zmienności danych wejściowych. Faza ta pozwala na dokonanie oceny stopnia „nauczenia” się analizowanych zależności przez sieć.

Tak przygotowana sieć neuronowa może zostać już wykorzystana dla celów uzyskania gotowej odpowiedzi jako reakcji na podany sygnał wejściowy.

Wykorzystując prawidłowo zbudowaną oraz „nauczoną” sieć, uzyskać można dobre wyniki przy jednocześnie krótkim czasie obliczeń i niskich wymaganiach sprzętowych.

Możliwe zastosowania

Jak już wspomniano wcześniej, sieci neuronowe znalazły szerokie zastosowanie w praktyce. Literatura podaje konkretne zastosowania, których przykłady zamieszczono poniżej [9]:

- NASA wykorzystuje sieci neuronowe do sterowania ramieniem manipulatora działającego w ładowni promów kosmicznych, co pozwala na utrzymywanie manipulowanych obiektów w niezmiennym położeniu w warunkach nieważkości,
- Uczni z New York University Medical Center zastosowali sieć jako alternatywę dla złożonych i czasochłonnych tensorowych obliczeń parametrów ruchu robota, co umożliwiło przyspieszenie działania systemu sterowania a przez to pracę robota w czasie rzeczywistym,
- Firma General Dynamics opracowała dla US Navy system oparty na sieci neuronowej klasyfikujący i rozpoznający sygnały sonarowe, pozwalający na identyfikację jednostki pływającej a nawet obiektów nadwodnych (np. helikopter unoszący się nad powierzchnią oceanu),
- Szpital Anderson Memorial Hospital w południowej Karolinie wykorzystał sieci neuronowe dla celów optymalizacji leczenia, pozwoliło na uzyskanie znacznych oszczędności a przede wszystkim na uratowanie życia kilkudziesięciu pacjentów,
- Producent raket Genera Devices Space Systems Division użył sieci neuronowych do sterowania pracą 150 zaworów doprowadzających paliwo i tlen do silników rakiety Atlas, co pozwoliło na zastąpienie dotychczas stosowanego kosztownego i zawodnego, złożonego systemu automatyki opartego na setkach sensorów,
- Firma Eaton Corporation wykorzystwała sieć neuronową w układzie sterowania wspomagającym pracę kierowcy dużej ciężarówki (pięć osi, osiemnaście kół) przy wykonywaniu niektórych szczególnie trudnych manewrów (np. cofanie z naczepą),
- Amerykańskie siły powietrzne (US Air Force) używają sieci neuronowe do rozwoju symulatorów lotu,
- Koncern Ford Motor Company przygotował nowy system diagnostyczny dla silników,
- Linie TWA stosują sieci neuronowe do lokalizacji bomb w swoim terminalu na lotnisku JFK w Nowym Yorku,

- W energetyce, np. w elektrowni BC Hydro w Vancouver sieci neuronowe znalazły zastosowanie do prognozowania zapotrzebowania na moc elektryczną,
- Firma Halliburton wykorzystuje sieci neuronowe do identyfikacji typu skał napotykanych podczas prowadzenia odwiertów przy poszukiwaniu złóż ropy i gazu.

Podsumowanie

W pracy poruszono problematykę sztucznych sieci neuronowych. Podano podstawowe zasady budowy i funkcjonowania sieci prezentując przy tym możliwe modyfikacje. Zaprezentowano przykłady istniejących konkretnych zastosowań sieci. Podano analogię z teorią systemów autonomicznych M. Mazura. Modele oraz systemy i układy funkcjonujące w oparciu o technikę sieci neuronowych stają się coraz częściej realną alternatywą dla dotychczas funkcjonujących rozwiązań. Działają one co prawda wyłącznie w środowisku wirtualnym, jednakże mogą być częścią korelatora systemów sterowanych.

Literatura

- [1] Nałęcz M., red. 2000, *Biocybernetyka i inżynieria biomedyczna 2000, tom 6: Sieci neuronowe*, Akademicka Oficyna Wydawnicza EXIT, Warszawa,
- [2] Hao Z., Kefa C., Jianbo M., 2001, *Combining neural network and genetic algorithms to optimize low NO_x pulverized coal combustion*, Fuel 2001; 80:2163-2168,
- [3] Kesgin U., 2003, *Genetic algorithm and artificial neural network for engine optimization of efficiency and NO_x emission*, Fuel 2004; 83:885-895,
- [4] Rutkowski L., red. 1996, *Sieci neuronowe i neurokomputery*, Wydawnictwo Politechniki Częstochowskiej, Częstochowa, Seria Monografie Nr 40,
- [5] Rutkowska D., Piliński M., Rutkowski L., 1997, *Sieci neuronowe, algorytmy genetyczne i systemy rozmyte*, PWN, Warszawa,
- [6] Kwater T., Kędzior Z., Twaróg B., 2001, *Estimation by artificial neural network in ecological problems*, AMSE-Conference MS'2001-Lviv (Ukraine) 23-26 May:212-215,
- [7] Chudzik S., Gryś S., Bąbka R., 2001, *Możliwość wykorzystania sztucznych sieci neuronowych do rozwiązania współczynnika zagadnienia odwrotnego*, Środkowoeuropejska IV Konferencja Naukowo-Techniczna: Metody i Systemy Komputerowe w Automatyce i Elektrotechnice-IV MSKAE 2001, Częstochowa-Poraj: 46-48,
- [8] Osowski S., 1996, *Sieci neuronowe w ujęciu algorytmicznym*, WNT, Warszawa,

- [9] Tadeusiewicz R., 1998, *Elementarne wprowadzenie do techniki sieci neuronowych z przykładowymi programami*, Akademicka Oficyna Wydawnicza PLJ, Warszawa,
- [10] Mazur M., 1999, *Cybernetyka i charakter*, Warszawa.